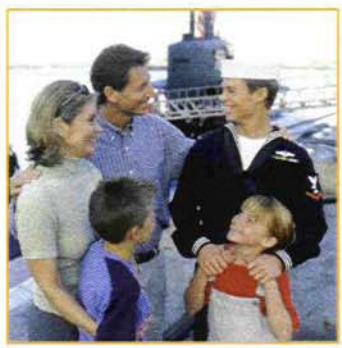








Weighting Report

2000 Military Recruiter Survey

August 2002

JAMRS Report No. 2002-004
August 2002

WEIGHTING REPORT FOR THE 2000 MILITARY RECRUITER SURVEY

Ismael Flores-Cervantes and Richard Valliant
Westat

Andrea B. Zucker
Defense Human Resources Activity

B. J. George
Defense Manpower Data Center

Department of Defense
Defense Human Resources Activity
Joint Advertising, Market Research and Studies
4040 N. Fairfax Drive, Arlington, VA 22203-1613

Table of Contents

Page

Table of Contents (continued)

Appendix

List of Tables

List of Figures

WEIGHTING REPORT FOR THE 2000 MILITARY RECRUITER SURVEY

Weighting Procedures for the 2000 Military Recruiter Survey

This report describes the weighting procedures for the analytical weights for the 2000 Military Recruiter Survey (MRS).

In order to produce estimates, weights are applied to sample data. In particular, sample weighting is carried out to accomplish the following objectives:

- Compensate for differential probabilities of selection
- Reduce biases occurring because nonrespondents may have different characteristics than respondents (differential response rates); and
- Improve the precision of the survey-based estimates

The analytical weights for the 2000 MRS were created in three steps. In the first step the base weights were computed as the inverse of the probability of selection of the sampled member. The sampled members were randomly drawn from a stratified frame without replacement. The sampling frame was compiled from lists of recruiters from the different services in the Armed Forces. The frame was stratified based on Service and region as shown in Table 1. In the second step of the weighting process, the base weights were adjusted to account for members whose eligibility to the survey could not be determined (members with unknown eligibility). These members neither returned a questionnaire nor provided any information to determine if the member had retired, separated from the military or was no longer recruiting. In the last step, the weights were adjusted for nonresponse among eligible members in the sample (eligible nonrespondents). These members were eligible but did not have usable survey data because each returned an incomplete questionnaire.

In other DoD surveys, there is an additional adjustment made to the weights. In the last step of the weighting, the weights are poststratified to control totals derived from updated frames. In the case of the MRS, there was no such frame available and the weights were not further adjusted. Postratification adjustments are also used to correct distortions in the sums of weights for some analytical variables caused by the nonresponse adjustments. In the case of the MRS, the sampling strata were used as nonresponse adjustment cells. Therefore, the sum of weights by stratum was preserved after the nonresponse adjustments.

Table 1.
Stratum Definition, Population Size and Sample Size for the 2000 MRS

Stratum	Service/ Region	Population Size	Sample Size
1	Air Force	1,952	1,431
2	Air Force Reserve	311	241
3	Air National Guard	471	285
4	Army National Guard	2,711	828
5	Army Region 1	1,700	523
6	Army Region 3	1,491	510
7	Army Region 4	1,470	508
8	Army Region 5	1,122	478
9	Army Region 6	1,610	518
10	Army Reserve	1,542	515
11	Coast Guard	382	262
12	Marine Corps Region East	1,211	743
13	Marine Corps Region West	1,428	877
14	Navy Area 1	1,111	477
15	Navy Area 3	1,272	493
16	Navy Area 5	1,303	497
17	Navy Area 8	1,384	503
18	Navy Area Other	59	17
19	Naval Reserve	724	420
	Total	23,254	10,126

Data files with the analytical weights were created so that variances of survey estimates can be computed using statistical packages such as SUDAAN®, SAS® or WesVar™. SUDAAN and SAS use the linearization method of variance estimation based on the Taylor series approximation while WesVar uses replication methods. Newer versions of SUDAAN (release 8 or later) can also compute variances using replication. SUDAAN was used to produce some estimates and their standard errors for the 2000 Military Recruiter Survey Overview Report while WesVar was used to compute the estimates for the 2000 Military Recruiter Survey Tabulation Volumes.

In the MRS there were few responding officers in the survey and including them in the public use file would risk disclosing their identities. Excluding the officers from the public use file would eliminate the possibility of users identifying sampled officers. Westat was asked to investigate the effect of excluding data for officer participants. The results of this analysis are presented in Appendix A. The findings show no significant difference for most estimates and their standard errors when officers are excluded

Assigning Disposition Codes for the 2000 Military Recruiter Survey

Each person in the 2000 Military Recruiter Survey (MRS) survey was assigned a disposition code indicating whether the person was an eligible respondent, an eligible nonrespondent, an ineligible, or a member whose eligibility status was unknown. These codes were a key input in weighting and in the computation of response rates, discussed in later sections. The final disposition code was assigned sequentially combining the information from the following two variables:

- RESULT—Survey Control System (SCS) disposition code assigned to each sampled member during the data collection; and
- COMPFLAG—Completed questionnaire indicator created during the weighting process.

The creation of these variables and the process for assigning the final disposition codes are described in the following sections. The eligibility for the survey was determined for all the recruiters in the sample. The recruiters whose eligibility status was known were classified as either eligible or ineligible members. Based on the return of a completed questionnaire, the eligible recruiters were classified as eligible respondents or eligible nonrespondents.

Survey Control System Disposition Code

The Survey Control System contained a variable with the survey disposition code (RESULT) as determined during the data collection period for each mailed survey. Sampled members were coded according to the type of return and/or any other information available during data collection. Returns were classified as nonblank questionnaires, blank questionnaires, final non-locatable members, ineligible members (members who retired, separated from the military or were no longer a recruiter) or other non-response. Table 2 shows the numbers of cases and descriptions for the values of the variable RESULT that appeared in the MRS sample.

Table 2.
Description of the Survey Control System Disposition Codes (RESULT) That Were Used in the 2000 MRS Recruiter

RESULT	Descriptions	Sample cases	% Sample cases	Sum of base weights	% Sum of base weights
01	Nonblank questionnaire	5,670	55.99	13,118	56.41
02	Returned blank	8	0.08	22	0.09
03	Final nonlocatable	481	4.75	1,085	4.67
06	Retired from military	20	0.20	42	0.18
07	Separated from military	6	0.06	13	0.06
08	No longer a recruiter	60	0.59	131	0.56
10	Other nonresponse	3,881	38.33	8,842	38.03
	Total	10,126	100.00	23,254	100.00

Completed Questionnaire

The variable that indicates whether a questionnaire was completed (COMPFLAG) was created using questions 2 (R00002) and 4 (R00004A to R00004E) from the questionnaire (Figure 1). A questionnaire was considered complete if the respondent answered both questions. Table 3 shows the distribution of COMPFLAG, the sums of base weights and the corresponding percentages in the MRS sample.

Figure 1.
2000 MRS Questions 2 and 4

R00002

2. How long have you been assigned to recruiting duty (include all tours in recruiting)?

① Less than one year
② 1 year, but less than 2
③ 2 years, but less than 3
④ 3 years, but less than 6
⑤ 6 or more years

R00004A – R00004E

3. Do you have specific monthly/annual goals/missions? (MARK ALL THAT APPLY)

○ Yes, personal monthly goals/missions
○ Yes, personal annual goals/missions
○ Yes, team monthly goals/missions
○ Yes team annual goals/missions
○ No, neither personal nor team goals/missions used

Table 3.
Complete Questionnaire Flag (COMPFLAG)

COMPFLAG	Sample cases	% Sample cases	Sum of base weights	% Sum of base weights
0 – Incomplete	4,487	44.31	10,203	43.88
1 – Complete	5,639	55.69	13,051	56.12
Grand Total	10,126	100.00	23,254	100.00

Final Disposition Codes

The method of assigning the final disposition codes was a sequential process that used the variables described in the previous sections. Once the disposition codes were assigned, each combination was checked for inconsistencies.

Table 4 lists the combinations of the variables RESULT and COMPFLAG that occurred in the MRS sample, the number of sampled cases, and the sums of base weights. Based on these two variables, a new variable denoted as ELIG was created with the following categories:

- *ER*— Eligible respondents. This group consisted of all eligible recruiters who participated in the survey and provided substantially complete and usable survey data.
- *ENR*— Eligible nonrespondents. This group consisted of all sampled recruiters who were known to be eligible for the survey, but did not provide complete (based on questions 2 and 4) and usable survey data.
- *IN*— Ineligibles or out-of-scope. This group consisted of members who had retired, separated from the military or who were no longer recruiters.
- *UNK*— Other nonrespondents whose eligibility was unknown. This group consisted of all the nonresponding persons for whose eligibility to the survey could not be determined. This group consisted of recruiters who did not return the questionnaire, postal non-deliveries and other non-locatable recruiters.

Table 4.
Combinations of Variables Used to Determine Disposition Codes

Row	Eligibility (ELIG)	Survey control system disposition code (RESULT)	Complete questionnaire (COMPFLAG)	Sampled cases	Sum of base weights
Eligible Respondents					
1	ER	01 Nonblank questionnaire	1	5,639	13,051
Eligible Nonrespondents					
3	ENR	01 Nonblank questionnaire	0	31	67
4	ENR	02 Returned blank	0	8	22
Ineligible as reported by self or proxy					
3	IN	06 Retired from military	0	20	42
4	IN	07 Separated from military	0	6	13
5	IN	08 No longer a recruiter	0	60	131
Unknown eligibility					
6	UNK	03 Final non-locatable	0	481	1,085
7	UNK	10 Other nonresponse, survey not returned	0	3,881	8,842
			Total	10,126	23,254

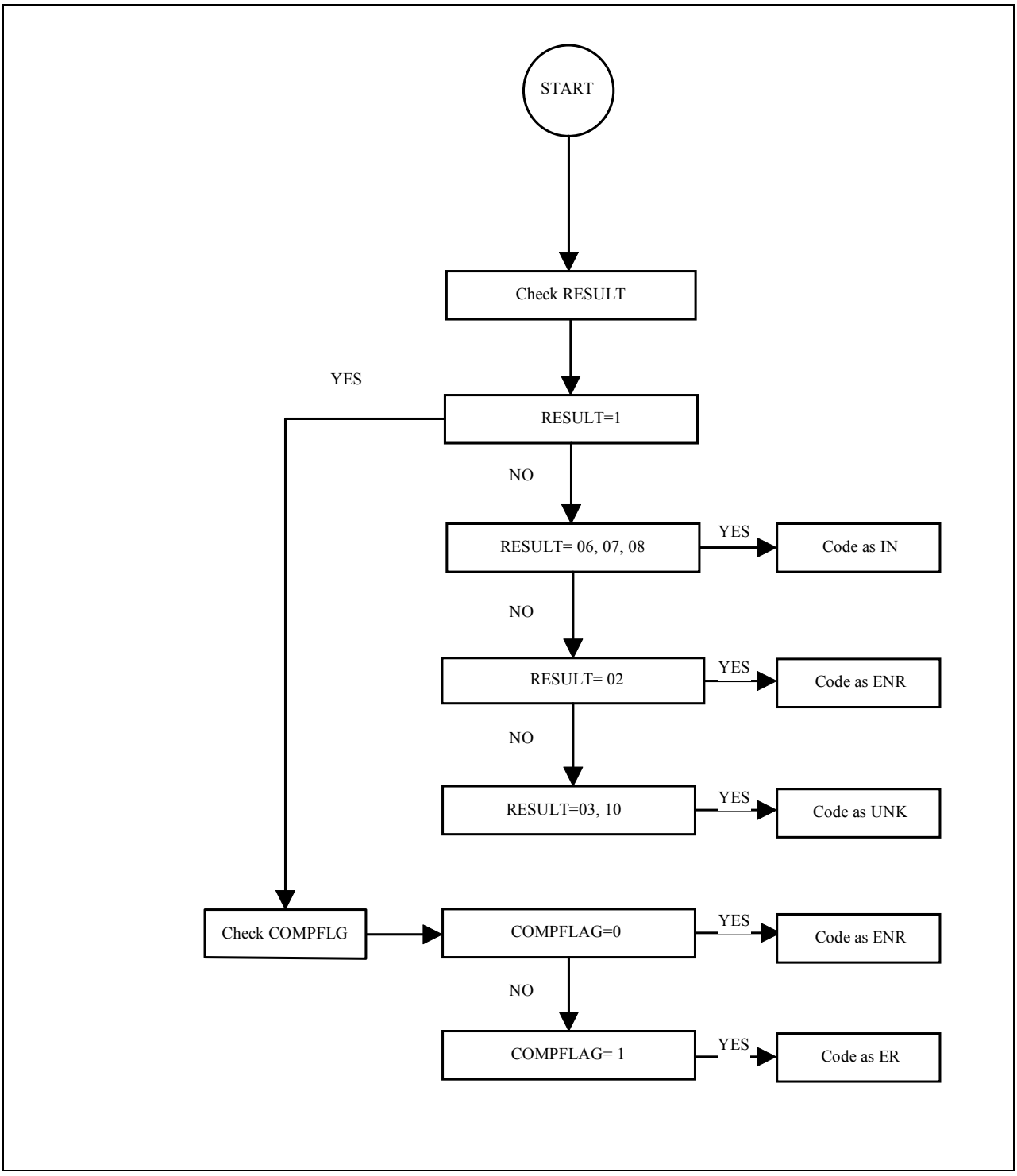
Table 5 lists the counts of cases, sums of base weights and percentages for each eligibility category.

Table 5.
Sampled Cases and Sums of Base Weights by Eligibility (ELIG)

ELIG		Sampled cases	% Sampled cases	Sum of base weights	% Sum of base weights
ER	(Eligible respondents)	5,639	55.7	13,051	56.1
ENR	(Eligible nonrespondents)	39	0.4	89	0.4
IN	(Ineligibles)	86	0.9	186	0.8
UNK	(Eligibility is unknown)	4,362	43.1	9,927	42.7
Grand Total		10,126	100.0	23,254	100.0

Figure 2 is a general flowchart showing how the disposition code ELIG was assigned. The Survey Control System (RESULT) code was used to divide the sample into groups for eligibles, ineligibles and members with unknown eligibility. The variable COMPFLAG was used to split the eligible members into eligible respondents (*ER*) and eligible nonrespondents (*ENR*) based on whether the questionnaire was complete or not.

Figure 2.
Sequential Assignment of ELIG Disposition Codes



Weighting Procedures

The analysis of survey data from complex sample designs requires the use of weights to (1) compensate for differential probabilities of selection; (2) adjust for differential response rates; and (3) improve the precision of the survey-based estimates (Skinner *et al*., 1989). To develop the weights for the 2000 MRS survey, the following steps were taken. First, base weights equal to the reciprocal of the probability of selection were assigned to each recruiter selected for the sample. Next, the base weights were adjusted for unknown eligibility and for nonresponse using weighting classes defined by the strata used in sample selection. Details of this weighting methodology are described in the following sections.

Calculation of Base Weights

The 2000 MRS sample was a stratified simple random sample selected without replacement. The overall probabilities of selection varied by design strata in order to satisfy the precision goals specified at the beginning of the study. Let U be the frame of the N units in the population (i.e., military recruiters at the time of sampling). Note that the frame size N included some members who were ineligible at the time the survey was conducted because, for example, they did not meet the criteria to be production recruiters. The frame U was partitioned into H non-overlapping strata $U_1,\ldots,U_H$ consisting of N_h units in each stratum h so that

$$N = \sum_{h=1}^{H} N_h.$$

A simple random sample of size n_h was selected without replacement within each stratum U_h. Given this design, the base weight for the i-th sampled recruiter in stratum h was calculated as:

$$w_{hi} = \frac{N_h}{n_h} \qquad i = 1,\ldots,n_h \cdot$$

For each individual classified in stratum h, the base weight was computed as the ratio of the total number of recruiters in the stratum to the stratum-level sample size. The base weight w_{hi}, equal to the reciprocal of the probability of selection, was attached to each sample unit in the data file. Note that n_h is the number of recruiters initially sampled in stratum h without regard to whether or not they ultimately participated in the survey.

Weighting Adjustments

In an ideal survey, all the units in the inference population are eligible to be selected into the sample; and all those that are selected participate in the survey. In practice, neither of these conditions usually occurs. Some of the sampled units do not respond (unit nonresponse); some sample units are discovered to be ineligible during the data collection period; and the eligibility status of some units cannot be determined. If these problems are not addressed, the estimates of the survey will be biased. We used nonresponse weight adjustments to deal with unknown eligibility and unit nonresponse. The following sections describe these methods in detail.

Nonresponse Adjustments

Unit nonresponse (i.e., whole questionnaire nonresponse) occurs when a sampled recruiter fails to respond for any reason. For example, nonresponse could result from failure to locate the recruiter because of mobility, incorrect addresses in the frame, or from the unwillingness to participate in the survey. Because the (unweighted) response rate (defined in a later section) in the 2000 MRS was substantially less than 100 percent, adjusting for unit nonresponse was an important step in attempting to reduce the bias of the estimates.

To compensate for losses due to nonresponse, the weights were adjusted in two stages: (1) The first stage of adjustment accounted for the fact that the eligibility status of some sampled members could not be determined. (2) The second stage of adjustment compensated for losses due to eligible sampled members who did not complete the questionnaire. At each stage the base weights of usable cases were inflated to account for ones that were unusable. These adjustments were done within classes that grouped persons with similar characteristics together.

This form of adjustment is referred to as sample weighting or weighting class adjustments since it adjusts the weighted distribution of the respondents across the weighting classes to that of the total sample (Kalton and Kasprzyk, 1986).

Nonresponse adjustment can increase the variability of the weights, and thus, tends to increase the sampling variance of some estimates (Kish, 1992). Ideally, the reduction in bias from using a nonresponse adjustment, more than compensates for the increase in variance. When the weighting class cells contain sufficient cases and the adjustment factors do not become either inordinately large or substantially different from each other, the effect on variances is modest. Very large adjustment factors or factors that are much different from others can occur in cells with high nonresponse rates or small numbers of respondents (i.e. less than 30 respondents in the cell). In the MRS, this situation was not an issue because the weight adjustments were done within design strata. With one exception, each stratum had a large number of respondents. In this case, the single small stratum (Stratum 17, Navy Area Other) was not combined with any other strata because each Navy stratum (Navy Areas 1, 3, 5, 8, and Other) was of separate analytic interest.

As discussed previously, each sampled recruiter was assigned to an appropriate response-status group (*ER*, *ENR*, *IN*, or *UNK*). At the first stage of weight adjustment, it was assumed that members with unknown eligibility (Group *UNK*) would have been distributed among the *ER*,

ENR, and *IN* categories had it been possible to determine their status. In this case, the first-stage nonresponse adjustment factor was calculated within stratum *h* as:

$$f_h^{A1} = \begin{cases} \dfrac{\sum_{i\in ER_h} w_{hi} + \sum_{i\in ENR_h} w_{hi} + \sum_{i\in IN\_SR_h} w_{hi} + \sum_{i\in UNK_h} w_{hi}}{\sum_{i\in ER_h} w_{hi} + \sum_{i\in ENR_h} w_{hi} + \sum_{i\in IN_h} w_{hi}} & \text{If the } i\text{-th sampled person classified in stratum } h \text{ belongs to response group } ER_h, ENR_h, \text{ or } IN_h. \\ 0 & \text{If the } i\text{-th sampled person in stratum } h \text{ is in } UNK_h. \end{cases}$$

The sums in the numerator of f_h^{A1} extend over the following types persons in stratum *h*: eligible respondents (*ER*), eligible nonrespondents (*ENR*), the ineligibles (*IN*), and the unknowns (*UNK*). The term w_{hi} is the base weight for the *i*-th sampled person in stratum *h*.

The first-stage nonresponse-adjusted weight w_{hi}^{A1}, for a sampled recruiter in stratum *h* was then computed as

$$w_{hi}^{A1} = f_h^{A1} w_{hi}$$

Thus, if persons with unknown eligibility accounted for 50 percent of the weight in stratum *h*, the weights on the other units would be increased by a factor of 2.

The second nonresponse adjustment increased the adjusted weight of eligible respondents to account for eligible nonrespondents. The second-stage nonresponse adjustment factor for stratum *h* was computed as:

$$f_h^{A2} = \begin{cases} \dfrac{\sum_{i\in ER_h} w_{hi}^{A1} + \sum_{i\in ENR_h} w_{hi}^{A1}}{\sum_{i\in ER_h} w_i^{A1}} & \text{If the } i\text{-th sampled person in stratum } h \text{ belongs to response group } ER_h. \\ 0 & \text{If the } I\text{-th sampled person sampled in stratum } h \text{ belongs to response group } ENR_h. \\ 1 & \text{If the } i\text{-th sampled person is in } IN_{\text{h}} . \end{cases}$$

The first sum in the numerator of f_h^{A2} for eligible respondents extends over the respondents (Group *ER*) in stratum *h;* the second over the eligible nonrespondents (Group *ENR*) in the stratum; and w_{hi}^{A1} is the previously adjusted weight of the *i*-th sampled recruiter.

The second-stage nonresponse-adjusted weight w_{hi}^{A2}, for the (*hi*)-th sampled recruiter was computed as:

$$w_{hi}^{A2} = f_h^{A2} w_{hi}^{A1}.$$

Thus, after the two stages of nonresponse adjustment, the weight for a respondent in stratum *h* can be written as

$$w_{hi}^{A2} = f_h^{A2} f_h^{A1} w_{hi}.$$

Note that after the two stages of nonresponse adjustments, the members with non-zero weights were those in *ER* and *IN* groups. The members with unknown eligibility (*UNK*) and eligible nonrespondents (*ENR*) have zero weight after the two adjustments.

Construction of Weighting Classes

The main objective in constructing weighting classes was to group respondents and nonrespondents with similar characteristics into cells. Ideally, the characteristics used for grouping should be related to both the likelihood of responding to the survey and to values of data items collected. Each of these characteristics must be available for all sampled persons.

For the MRS, the sampling strata were used as weighting classes. There are 19 strata created using Service and region listed in Table 6. Table 6 also shows the adjustment factors for unknown eligibility and nonresponse.

Table 6.
2000 MRS Nonresponse Adjustment Cells and Factors

Stratum	Service/ Region	Adjustment for unknown eligibility f_h^{A1}	Adjustment for eligible nonrespondents f_h^{A2}	Overall adjustment factor $f_h^{A1} f_h^{A2}$
1	Air Force	1.884	1.003	1.889
2	Air Force Reserve	2.104	1.027	2.162
3	Air National Guard	1.647	1.012	1.668
4	Army National Guard	1.595	1.008	1.608
5	Army Region 1	1.861	1.004	1.868
6	Army Region 3	1.962	1.016	1.992
7	Army Region 4	1.795	1.004	1.801
8	Army Region 5	1.707	1.004	1.713
9	Army Region 6	1.655	1.006	1.666
10	Army Reserve	1.951	1.008	1.966
11	Coast Guard	1.432	1.011	1.448
12	Marine Corps Reg. East	1.867	1.005	1.876
13	Marine Corps Reg. West	1.683	1.010	1.700
14	Navy Area 1	1.651	1.011	1.668
15	Navy Area 3	1.700	1.007	1.712
16	Navy Area 5	1.630	1.000	1.630
17	Navy Area 8	1.548	1.006	1.557
18	Navy Area Other	2.833	1.000	2.833
19	Naval Reserve	2.100	1.005	2.111

Computation of Variance for Estimates for the 2000 MRS

Variance estimation procedures have been developed to account for complex sample designs such as the selection of a sample in multiple stages and the use of differential sampling rates to oversample a targeted subpopulation. The two main methods for estimating variances from a complex survey are linearization using the Taylor series approximation (theory-based) and replication (empirical). Wolter (1985) is a useful reference on the theory and applications of these methods. Shao (1996) is a more recent review paper that compares the methods. The next two sections describe how these methods were implemented to compute variances of the estimates for the 2000 MRS survey. An in depth discussion of software applications for analysis of the 2000 MRS, complete with examples, can be found in the *2000 MRS Administration, Data Sets and Code Book*, Appendix J.

Taylor Series Method to Compute Variances

A widely used method for estimating variances in complex surveys is based on the Taylor series approximation. A linear approximation to a statistic is formed from the Taylor series

expansion for the function of interest. This approximation is then substituted into the variance formula appropriate for the sample design. The Taylor series method relies on the simplicity associated with estimating the variance for a linear statistic even with a complex sample design and is valid in large samples. In this formulation, the variance strata and primary sampling units (PSUs) must be defined.

SUDAAN® (Software for the Statistical Analysis of Correlated Data) (Research Triangle Institute 2001) is a software package designed to produce variance estimates for complex surveys using the Taylor series method. SUDAAN computes standard errors of the estimates taking into account most features of complex sample designs and estimators. SUDAAN is also capable of reflecting stratum-by-stratum finite population correction (*fpc*) factors in the computation of variances. This is particularly important for the 2000 MRS survey, where some strata are sampled at high rates. Recent releases of SUDAAN (Release 8 and later) can also compute estimates of variance based on replication methods.

For descriptive statistics, SUDAAN offers three procedures: PROC CROSSTAB for categorical variables, PROC DESCRIPT for continuous variables and PROC RATIO for ratios of totals. These procedures can be used to compute statistics of interest, such as estimated totals, means, and percentages along with their corresponding standard errors, design effects, and confidence intervals. SUDAAN can be used to reflect the facts that:

(i) The frame contains recruiters who self-reported or were proxy-reported as ineligible, or would had been found ineligible had they been surveyed, and

(ii) The *fpc* is important in some strata.

Differences of table cell estimates can also be computed in PROC DESCRIPT and PROC RATIO. The statements that control these calculations are CONTRAST, DIFFVAR, and PAIRWISE.

To reflect the effect of the design in variance estimation, SUDAAN requires variables that identify the variance estimation strata and primary sampling units (PSUs). The variance estimation strata are the original design strata from which the sample was drawn. For the 2000 MRS the variance estimation strata were the service and region strata and , the sampled PSU corresponds to the individual sampled person. It should be noted that small sample sizes could lead to unstable variance estimates. Normally this problem is solved by collapsing original strata with fewer than 30 respondents; however, for one design stratum, Navy Area Other, this approach was not used. The Navy Area Other stratum was not collapsed with other strata because it needed to be analyzed separately.

The variance strata and PSU-identifying variables were part of the data set delivered to DMDC so estimates and their standard errors can be computed using SUDAAN.

SAS (version 8 and later) has two procedures for analyzing survey data: PROC SURVEYMEANS and PROC SURVEYREG. Both use the Taylor Series linearization approach to estimate standard errors. SURVEYMEANS produces estimates of means, proportions, and

totals along with their corresponding standard errors, while SURVEYREG fits linear regression models (logistic regression is not yet available). No design effects are estimated with either PROC. Estimates of differences or other linear combinations are not available in SURVEYMEANS.

These procedures are new in SAS and do not contain as many features as some other packages. Finite population correction factors can be included in variance estimates for MRS but the effect of nonresponse adjustments cannot.

Replication Methods

The basic idea behind replication is to draw subsamples from the full sample, compute the estimate from each of the subsamples, and estimate the variance from the subsample estimates. The subsamples are called replicates and the estimates from the subsamples are called replicate estimates. Rust and Rao (1996) discuss replication methods, show how the units included in the subsamples can be defined using variance strata and units, and describe how these methods can be implemented using weights.

Replicate weights are created to generate a corresponding set of replicate estimates. Each replicate weight is constructed using the same estimation steps as the full sample weight, but using only the subsample of cases composing each replicate. Once the replicate weights are developed, it is straightforward to compute estimates of variance for sample estimates of interest.

WesVar™ (Westat, 2000) is a computer software program that generates measures of variability (e.g., standard errors, coefficients of variation, and confidence intervals) for estimates using a specified set of replicate weights. WesVar allows derived statistics, like differences or ratios, to be calculated using the Cell Function feature of tables.

Using replication to estimate variances reflects the effects of the design and the nonresponse adjustments. Also included are provisions to approximately reflect the finite population correction factors in the computation of variances. When using WesVar, no extra statements are needed for variance estimation for subgroups of interest and, therefore, no knowledge of the sample design is required.

For reference, Table 7 lists some of the features available in SUDAAN, SAS, and WesVar that are relevant to MRS analysis. This list is not exhaustive, particularly for SUDAAN and WesVar which include, other analysis features in SUDAAN and WesVar that may also be of interest to data users.

Table 7.
Features of Three Software Packages for the Analysis of Survey Data

Feature	SUDAAN	SAS	WesVar
Estimation features reflected in variance estimates			
Stratification	x	x	x
Ineligible cases in poststratification frame	x	x	x
Differential weights among cases	x	x	x
Nonresponse adjustments (unknown eligibility, eligible nonrespondents)	x*	NA	x
Poststratification	x	NA	x
Finite population correction factors	x	x	x **
Tables			
Totals/standard errors	x	x	x
Means/standard errors	x	x	x
Proportions/standard errors	x	x	x
Multi-way tables	x	x	x
Differences of cell estimates/standard errors	x	NA	x
Ratios of cell estimates	x	NA	x
Linear regression			
Parameter estimates/standard errors	x	x	x
Confidence intervals for parameters	x	x	x
Logistic regression			
Parameter estimates/standard errors	x	NA	x
Confidence intervals for parameters	x	NA	x
Odds ratios/confidence intervals	x	NA	x
Multinomial logistic regression (unordered categories)			
Parameter estimates/standard errors	x	NA	x
Odds ratios/confidence intervals	x	NA	x
Multinomial logistic regression (ordered categories)			
Parameter estimates/standard errors	x	NA	NA
Odds ratios/confidence intervals	x	NA	NA

Note: NA= not available.
* Available in SUDAAN when estimates based on replication methods are computed.
**Common fpc's at the replicate level

The Jackknife Method

The method of replication used for the 2000 MRS is known as the stratified, delete-one jackknife. The general procedure is to form groups of sampled persons, and then to form replicates or subsamples by deleting one group at a time. The method is called JKn in WesVar. The method is discussed in some depth in Chapter 4 of Wolter (1985) and in Rust (1986).

To implement the method, variance strata (denoted in WesVar as *VARSTRAT*) and variance units (denoted as *VARUNIT*) were created. The variance strata are combinations of

design strata. The variance units are groups of initial sampled persons, including eligibles, ineligibles, and unknowns. Let $\tilde{h}$ be a variance stratum and denote the number of *VARUNIT*s in stratum $\tilde{h}$ by $n_{\tilde{h}}$. Since one *VARUNIT* is omitted at a time in the JKn method, the total number of replicate estimates is

$$G = \sum_{\tilde{h}=1}^{\tilde{H}} n_{\tilde{h}}$$

where $\tilde{H}$ is the number of variance strata. Note that $\tilde{H}$ may be different from the number of design strata.

Let *g* denote a particular combination of *VARSTRAT* and *VARUNIT*. Denote the replicate estimate formed by deleting *g* by $\hat{\theta}_{(g)}$. Because one *VARUNIT* is omitted at a time for JKn, *g* can be used to identify the *VARUNIT* itself, the set of sampled units (i.e., the replicate) that remains after omitting unit *g*, and the estimate computed from that replicate set of sampled units.

The weights used in calculating $\hat{\theta}_{(g)}$ account for the deletion of *g* from the sample as follows. Suppose that *g* identifies a *VARUNIT* in *VARSTRAT* $\tilde{h}$. When *g* is omitted, the base weights associated with the other $n_{\tilde{h}} - 1$ variance units in *VARSTRAT* $\tilde{h}$ are multiplied by the factor:

$$\frac{n_{\tilde{h}}}{n_{\tilde{h}} - 1}.$$

The base weight for *g* is multiplied by 0. The weights on all *VARUNIT*s in all other *VARSTAT* are unchanged. The two nonresponse adjustment steps described above, are then carried through using the sampled units in replicate *g* and their modified base weights. The estimate from replicate *g*, $\hat{\theta}_{(g)}$, thus, reflects all stages of weighting.

The JKn variance estimate for the full sample estimate $\hat{\theta}$ is then

$$v(\hat{\theta}) = \sum_{g=1}^{G} f_g h_g \left[\hat{\theta}_{(g)} - \hat{\theta}\right]^2$$

where f_g is the finite population correction (*fpc*) factor associated with the variance stratum containing unit *g* and $h_g = \left(n_{\tilde{h}} - 1\right) / n_{\tilde{h}}$ where $\tilde{h}$ is the *VARSTRAT* that contains unit *g*. The h_g are referred to as "JKn factors." In forming variance strata, it is important to put design strata

having the same or nearly the same *fpc* together in a variance stratum. This can be done only approximately since the sampling rates vary considerably among the MRS design strata.

Each sampled person's record in the data file have $G+1$ weights attached—one for the full sample and *G* replicate sample weights, computed as described above. In WesVar a data set called a *VAR* file is created that contains an indicator that the JKn method was used to create weights, the weights themselves, the finite population correction factors, and the h_g factors. When a user does tabulations or other analyses in WesVar using the *VAR* file, WesVar automatically evaluates variances using the JKn formula. The elaborate steps involved in creation of the weights and their proper usage are transparent to the user.

Number of Replicates

A key step in designing the replicate structure is to determine the number of replicates. The choice of the number of replicates is based on the desire to obtain adequate degrees of freedom (*DF*) to ensure stable estimates of variance while not having so many as to make the time or cost of computing variance estimates unnecessarily high. At *DF*=30, percentiles of the *t*-distribution are near those for the normal distribution; at *DF*=60, they are virtually the same as those for the normal. A rule of thumb is, thus, that at least 30 degrees of freedom are needed to obtain relatively stable variance estimates. The stability of a variance estimate for a subgroup is related to the number of *VARSTRAT* and *VARUNIT*s contributing to the subgroup estimate.

Note that having adequate *DF* is not a concern in SUDAAN because the linearization variance estimates have thousands of degrees of freedom for full sample estimates. Domain estimates have variances with fewer *DF* but enough to insure stability for most domains.

Formation of Replicates

Ideally, the creation of the replicate should be restricted to include the records from a single stratum only. Under this ideal approach, it is possible to correctly reflect the effect of the *fpc* in that specific stratum when JKn replicates are used. Note that the inclusion of the *fpc* (factor f_g) is only possible at the replicate level. At the same time, as described above, at least 30 replicates per stratum need to be created for better estimates at the stratum level. Then the total number of replicates to create would be approximated as

$$\text{Total replicates} \geq 30 * (\text{Number of strata})$$

The 2000 MRS survey has 19 strata, and with the rule above the required number of replicates needed to fully reflect the *fpc* in each design stratum would be about 570. Such a large number of replicates would be burdensome in practice. To solve this problem, we used an overall *fpc* for groups with similar sampling fractions and collapsed design strata when the variance strata were created. The *fpc* for a stratum *h* is

$$fpc_h = 1 - r_h^* = 1 - \frac{n_h^*}{N_h}$$

where

r_h^* = the achieved sampling fraction or sampling rate defined as the ratio of the achived sample size n_h^* (i.e., the number of respondents) divided by the population size N_h in the stratum.

Four zones of strata were created such that the design strata within a zone all had approximately the same *fpc*. The zones were then equated to the *VARSTRAT* for use in WesVar. Table 8 shows the ranges of stratum sampling rates in each zone and the number of design strata in each.

Table 8.
Replicate Zones for the 2000 MRS

Zone	Range of sampling rate	Number of strata	Percent of population
1	(0.384, 1]	3	46
2	(0.266, 0.384]	4	27
3	(0.205, 0.266]	5	16
4	(0, 0.205]	7	11
Total		19	100

An overall *fpc* factor is applied to the strata within each zone. The overall *fpc* factor is computed using the minimum sampling rate within the zone. Using the minimum group rate within the zone computes an actual stratum *fpc* for the zone with the smallest sample size and an approximation of the actual stratum *fpc* for the remaining zones. The overall *fpc* is larger than the actual stratum *fpc* leading to an overestimation of the variance for estimates for these strata. As a result, this procedure yields somewhat conservative (overestimated) variance estimates. Nevertheless, substantial improvements are expected in the precision of some domain estimates compared to the case where the *fpc* is ignored entirely. The *fpc*s for each zone for the 2000 MRS are shown in Table 9.

Table 9.
Overall fpc for the Replicate Zones

Zone	Minimum sampling rate	Overall *fpc* factor
1	0.4019	0.5981
2	0.3191	0.6809
3	0.2437	0.7563
4	0.1017	0.8983

An alternative is to use an average-based *fpc* computed using the average of the sampling rates of the strata within each zone. However, in this case, the variance can be underestimated for all the strata with a *fpc* larger than the average-based *fpc*.

To reduce the number of replicates, the design strata can be collapsed (or "folded") into pseudo-strata or variance strata (*VARSTRAT*). The number of variance strata and the number of replicates created within each variance stratum affect the number of degrees of freedom of the estimate of variance. As described before, each design stratum should ideally contain at least 30 replicates. For simplicity, the replicate zones were used as variance strata for the MRS. Table 10 shows the number of variance strata and number of replicates created within each variance stratum.

Table 10.
VARSTRAT and VARUNIT for the 2000 MRS

VARSTRAT	Number of replicates(VARUNIT)	JKn factor(h_g)
1	45	.977778
2	45	.977778
3	45	.977778
4	45	.977778
Total	180	

To assign the value of *VARUNIT*, all the records were sorted in the same random order in which they were sampled within *VARSTRAT*. The value of *VARUNIT* was a sequential number starting from 1 that was assigned to each record. When the sequential number reached the maximum number of *VARUNIT* within *VARSTRAT*, it restarted at one. This process was repeated until each recruiter record was assigned a *VARUNIT*. For example, in *VARSTRAT*=1 (i.e., zone =1) the records were serially numbered 1, 2, …, 45; 1, 2, …,45; and so on. All of the records numbered 1 were assigned to *VARUNIT* 1; all of the records numbered 2 were assigned to *VARUNIT* 2, and so on. The records with *VARUNIT*=1 were, thus, a subsample of the sample from all design strata assigned to *VARSTRAT*=1, as were the records in the other *VARUNIT*s. Because the ordering of the sampled persons was random, this method effectively divides the sample into random groups for each *VARSTRAT*.

To create the replicates, a series of factors REPF $\left(\tilde{h}, g\right)$ (replicate factor for *VARUNIT*=g in *VARSTRAT*= $\tilde{h}$) were created with the following values:

$$REPF\left(\tilde{h}, g\right) = \begin{cases} 0 & \text{if the person is in VARSTRAT} = \tilde{h} \text{ and VARUNIT} = g \\ \dfrac{n_{\tilde{h}}}{n_{\tilde{h}} - 1} & \text{if the person is in VARSTRAT} = \tilde{h} \text{ and VARUNIT} \neq g \\ 1 & \text{if the person is in VARSTRAT} \neq \tilde{h} \end{cases}$$

where $n_{\tilde{h}}$ = the number of *VARUNIT*s in *VARSTRAT* = $\tilde{h}$

The replicate base weight is the product of *REPF* $\left(\tilde{h}, g\right)$ and the full-sample base weight.

Calculation of Response Rates

Several rates for the 2000 MRS were computed in accordance with the standards defined by the Council of American Survey Research Organizations (CASRO, 1982). The rates are referred to as:

- Location rate (LR)
- Completion rate (CR)
- Response rate (RR)

These quantities were computed in such a way that *RR* = *LR* * *CR*. The rates are adjusted, as described below, to account for the fact that the eligibility of some units is unknown.

The *location rate* used for the 2000 MRS is

$$LR = \frac{\text{adjusted located sample}}{\text{adjusted eligible sample}} = \frac{N_L}{N_E}.$$

The *completion rate* is defined as

$$CR = \frac{\text{usable responses}}{\text{adjusted located sample}} = \frac{N_R}{N_L}.$$

The *response rate* is defined as

$$RR = \frac{\text{usable responses}}{\text{adjusted eligible sample}} = \frac{N_R}{N_E}$$

with $N_L, N_E,$ and N_R, defined below.

The adjustments account for the fact that the eligibility status of some persons is unknown so that the proportion of eligibles among the unknowns must be estimated. To facilitate computation of the CASRO rates, a separate code (CAS_ELIG) was created that identifies cases that contribute to the components of *LR*, *CR*, and *RR*, as defined in Table 11.

Table 11.
Disposition Codes for CASRO Response Rates (CAS_ELIG)

Eligibility code for CASRO response rates *(CAS_ELIG)*	**Weighting eligibility code *(ELIG)***	**Description**
ER	*ER*	Eligible respondent (usable)
ENR_NOQCOMP	*ENR*	Eligible nonrespondent (questionnaire not completed or returned blank questionnaire)
IN	*IN*	Ineligible
UNK_NOLOC	*UNK*	Unknown eligibility (nonlocatable recruiter)
UNK_OTHER	*UNK*	Unknown eligibility (questionnaire not returned)

The expressions for the numbers of located persons, eligible persons, and usable responses in terms of CAS_ELIG are given below. As a notational shorthand, CAS_ELIG codes are used to stand for counts of persons in the formulas. For example, ER denotes the count of eligible respondents.

$N_L =$ (Eligible respondents) + (Eligible nonrespondents) + (Estimate of eligible among unknowns who were located but did not return a questionnaire)

$$= ER + ENR + UNK\_OTHER\left(\frac{ER + ENR}{ER + ENR + IN}\right)$$

$$= ER + ENR + UNK\_OTHER\left(P\_E\right)$$

where $P\_E = \frac{ER + ENR}{ER + ENR + IN}$.

$N_E =$ (Eligible respondents) + (Estimate of eligible among unknowns)

$$= ER + ENR + (UNK\_OTHER + UNK\_NOLOC)\left(\frac{ER + ENR}{ER + ENR + IN}\right)$$

$$= ER + ENR + UNK(P\_E)$$

where $UNK = UNK\_OTHER + UNK\_NOLOC$.

$$N_R = (\text{Usable responses})$$

$$= ER$$

The adjusted located count, N_L, and the adjusted eligible count, N_E, can also be expressed by subtracting various counts from the total sample as shown below.

$N_E =$ Adjusted eligible sample

= (Total sample)

-(Known ineligibles)

-(Estimate of ineligibles among non-located unknwons)

-(Estimate of ineligibles among other unknowns)

$$= TOTAL - IN - (UNK\_NOLOC + UNK\_OTHER)\left(\frac{IN}{ER + ENR + IN}\right)$$

$$= ER + ENR + UNK\ (P\_E).$$

using the facts that $TOTAL = ER + ENR + IN + UNK\_NOLOC + UNK\_OTHER$ and $IN/(ER + ENR + IN) = 1 - P\_E$.

N_L = Adjusted located sample

= (Total sample)

– (Known ineligibles)

– (Non-located unknowns)

– (Estimate of ineligibles among other unknowns)

$$= TOTAL - IN - UNK\_NOLOC - UNK\_OTHER\left(\frac{IN}{ER + ENR + IN}\right)$$

$$= ER + ENR + UNK\_OTHER(P\_E)$$

Both weighted and unweighted location, completion, and response rates were calculated for the strata used in the sample design and are shown in Table 12. Weighted and unweighted rates are also reported for the full sample, and summary rates for Services and strata were computed. In all cases base weights were used in computing the weighted rates.

Table 12.
Unweighted and Weighted Location, Completion, and Response Rates for the Full Sample and Categories of Service and Stratum

Group	Adjusted eligible sample	Adjusted located sample	Complete responses	Unweighted Location rate	Unweighted Completion rate	Unweighted Response rate	Weighted Location rate	Weighted Completion rate	Weighted Response rate
Full sample	9,974	9,500	5,639	95.3%	59.4%	56.5%	95.3%	59.7%	56.9%
Service									
Army Active	2,532	2,413	1,405	95.3%	58.2%	55.5%	95.3%	58.1%	55.4%
Army National Guard	805	779	501	96.7%	64.3%	62.2%	96.7%	64.3%	62.2%
Army Reserve	497	465	253	93.6%	54.4%	50.9%	93.6%	54.4%	50.9%
Air Force Active	1,394	1,302	738	93.4%	56.7%	52.9%	93.4%	56.7%	53.0%
Air National Guard	268	264	161	98.6%	60.9%	60.0%	98.6%	60.9%	60.1%
Air Force Reserve	238	223	110	93.8%	49.3%	46.3%	93.8%	49.3%	46.3%
Marine Corps	1,604	1,541	903	96.0%	58.6%	56.3%	96.0%	58.6%	56.3%
Navy Active	1,962	1,875	1,193	95.6%	63.6%	60.8%	95.5%	63.7%	60.8%
Naval Reserve	418	385	198	92.1%	51.4%	47.4%	92.1%	51.4%	47.4%
Coast Guard	256	252	177	98.5%	70.1%	69.1%	98.5%	70.2%	69.1%
Sampling Stratum									
Air Force	1,394	1,302	738	93.4%	56.7%	52.9%	93.4%	56.7%	53.0%
Air Force Reserve	238	223	110	93.8%	49.3%	46.3%	93.8%	49.3%	46.3%
Air National Guard	268	264	161	98.6%	60.9%	60.0%	98.6%	60.9%	60.1%
Army National Guard	805	779	501	96.7%	64.3%	62.2%	96.7%	64.3%	62.2%
Army Region 1	521	491	279	94.3%	56.8%	53.5%	94.3%	56.8%	53.5%
Army Region 3	510	476	256	93.3%	53.8%	50.2%	93.3%	53.8%	50.2%
Army Region 4	508	488	282	96.1%	57.8%	55.5%	96.1%	57.8%	55.5%
Army Region 5	475	453	277	95.4%	61.2%	58.4%	95.4%	61.2%	58.4%
Army Region 6	518	505	311	97.5%	61.6%	60.0%	97.5%	61.6%	60.0%
Army Reserve	497	465	253	93.6%	54.4%	50.9%	93.6%	54.4%	50.9%
Coast Guard	256	252	177	98.5%	70.1%	69.1%	98.5%	70.2%	69.1%
Marine Corps Reg. East	734	706	391	96.2%	55.4%	53.3%	96.2%	55.4%	53.3%
Marine Corps Reg. West	870	835	512	95.9%	61.4%	58.8%	95.9%	61.4%	58.8%
Navy Area 1	470	461	282	97.9%	61.2%	60.0%	97.9%	61.2%	60.0%
Navy Area 3	488	463	285	94.9%	61.5%	58.4%	94.9%	61.5%	58.4%
Navy Area 5	487	458	299	94.0%	65.3%	61.4%	94.0%	65.3%	61.4%
Navy Area 8	500	477	321	95.4%	67.3%	64.2%	95.4%	67.3%	64.2%
Navy Area Other	17	17	6	100.0%	35.3%	35.3%	100.0%	35.3%	35.3%
Naval Reserve	418	385	198	92.1%	51.4%	47.4%	92.1%	51.4%	47.4%

References

Council of American Survey Research Organizations (1982). *On the definition of response rates* (special report of the CASRO task force on completion rates, Lester R. Frankel, Chair). Port Jefferson, NY: Author.

Kalton, G. and Kasprzyk, D. (1986). The Treatment of Missing Survey Data. *Survey Methodology,* 12, 1–16.

Kish, L. (1992). Weighting for Unequal Pi. *Journal of Official Statistics,* 8, 183–200.

Rust, K.F. (1986). Efficient Replicated Variance Estimation. *1986 Proceedings of the Section on Survey Research Methods* (pp. 81-87). Alexandria VA: American Statistical Association.

Rust, K.F. and Rao, J. N. K. (1996). Variance Estimation for Complex Surveys Using Replication Techniques. *Statistical Methods in Medical Research*, 5: 282–310.

SAS® System for Windows (Release 8.02) [Computer Software]. (2001). Cary, NC: SAS Institute Inc.

Skinner, C., Holt, D., and Smith, T., eds. (1989). *Analysis of Complex Surveys*. New York: J. Wiley & Sons.

Shao, J. (1996). Resampling Methods in Sample Surveys, (with Discussion). *Statistics*, 27, 203-254.

SUDAAN®, (Release 8.0) [Computer Software]. (2001). Research Triangle Park: Research Triangle Institute

WesVar™ (Version 4.0) [Computer software]. (2000). Rockville, MD:Westat.

Wolter, K. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.

APPENDIX A

Officers in 2000 MRS

Officers in the 2000 MRS

At the start of the MRS 2000 weighting process, Westat was asked to investigate the effect of excluding data for officer participants from the public use file. There were few responding officers in the survey and including them in the public use file would risk disclosing their identities.

To investigate this, Westat proposed calculating the weights (including officers) and then using SUDAAN to estimate standard errors for a few items. Two approaches were suggested for the SUDAAN estimates: 1) compute the estimates with enlisted participants treated as a domain and 2) compute the estimates on a file containing enlisted participants only. This appendix presents the comparison of the two approaches.

Since there are so few officer participants in the 2000 MRS, concern about confidentiality is legitimate. Excluding the officers would eliminate the possibility of a public user identifying an officer on the file. Officer participants comprise less than one percent of the total number of respondents; their weighted contribution is so small that excluding them would not change the weighted totals significantly. The primary issue is whether accurate standard errors can be computed from a public use data set that excludes officers. When a standard error is estimated for the domain of enlisted persons, officers should be assigned zero data values for the computation. If the estimate refers to a table cell or domain containing few, if any, officers, then the effect on the standard error of using a file with no officers should be small.

However, the remaining enlisted personnel should be treated as a domain when making estimates that exclude the officers. Treating the enlisted as a domain will increase standard errors compared to treating the sample as if it contained no officers. On the other hand, if officers were omitted from the public use file, users would not be able to make standard error calculations that were appropriate. But, since there were few sample officers, this might be a minor issue.

As with other DMDC surveys, Westat typically runs SUDAAN and WesVar on a select number of items of importance. This is done to check for reasonableness in the weighted estimates. A similar set of items was reviewed for the 2000 MRS survey. In addition, enlisted and officer participants were identified to create a domain variable.

First, the new variable, ENLISTED, was created to identify enlisted and officer participants; this variable was derived from the self-reported pay grade in the questionnaire. Table A-1 lists the number of responding sampled officers by stratum. The sample has 23 officers, 21 of which are in stratum 19, Naval Reserve.

Next, two files were created; the first file contained records of all participants and the second file contained records of enlisted participants only. The first file was used to produce SUDAAN and WesVar standard errors by domain. Likewise, the second file was used to produce SUDAAN standard errors for enlisted participants only. Table A-2 shows the standard errors from each run. In addition, the table provides two sets of ratios for each item; the median of the standard error for each item; the overall median across all items; and the median of the

totals across all items. The SUDAAN column labeled "Enlisted as a domain" gives the standard errors that are appropriately computed by setting the officers values to zero.

In summary, the findings show no significant difference in the treatment of enlisted participants for most estimates. The SUDAAN standard errors for each item are almost identical under the two approaches. All of the ratios are one or close to one with the exception of two items. The ratio for the Naval Reserve category of (self-reported) Service is slightly lower than the other categories (0.843). The ratio is substantially less than 1 for the Naval Reserve category of Stratum (0.230). This is because 21 of the 23 responding officers are in the Naval Reserve stratum and SUDAAN zeroes out their data values under the domain approach. Using the file of enlisted participants only, zeroes for the officers cannot be incorporated in the standard error calculation, resulting in a much smaller standard error. The ratios for the marginals of each variable are less than one since all officers are in the marginal. (Note that estimated totals differ among the variables due to different amounts of missing data.) Overall, excluding officers from the public use file would have very little effect on most standard errors.

For comparison, we also include WesVar standard errors and the ratio of the WesVar SEs to the SUDAAN domain SEs. The median ratio is 1.038 across all item categories and 0.997 across the totals. Thus, the WesVar standard errors were slightly larger than the proper SUDAAN standard errors for most estimates.

Table A-1.
Distribution of Sampled Officers by Stratum

Stratum	Item	Number of sampled officers
1	Air Force	0
2	Air Force Reserve	0
3	Air National Guard	0
4	Army National Guard	1
5	Army Region 1	0
6	Army Region 3	0
7	Army Region 4	0
8	Army Region 5	0
9	Army Region 6	1
10	Army Reserve	0
11	Coast Guard	0
12	Marine Corps RegionEast	0
13	Marine Corps Region West	0
14	Navy Area 1	0
15	Navy Area 3	0
16	Navy Area 5	0
17	Navy Area 8	0
18	Navy Area Other	0
19	Naval Reserve	21
Total	Total	23

Table A-2.
A Comparison of Estimates Made From a File with Enlisted Treated as a Domain and a File Containing Enlisted Members Only

Item	Estimated Total	SUDAAN SEs		WesVar SEs	Ratio	
		Enlisted as a domain	Enlisted members only	Enlisted as a domain	SUDAAN enlisted only/ SUDAAN enlisted as a domain	WesVar domain /SUDAAN enlisted as a domain
Satisfaction: Recruiting						
Very satisfied	2,193.158	78.275	78.307	84.271	1.000	1.077
Satisfied	7,259.710	128.562	128.444	136.493	0.999	1.062
Neither satisfied nor dissatisfied	4,443.289	110.941	110.963	109.402	1.000	0.986
Dissatisfied	4,885.591	116.024	116.043	120.739	1.000	1.041
Very dissatisfied	3,943.553	105.805	105.800	103.658	1.000	0.980
Total	22,725.301	40.887	38.090	43.050	0.932	1.053
Median					1.000	1.041
Satisfaction: Military Life						
Very satisfied	9,442.537	137.624	137.601	157.997	1.000	1.148
Satisfied	9,982.788	139.554	139.417	152.239	0.999	1.091
Neither satisfied nor dissatisfied	1,932.516	79.495	79.522	86.817	1.000	1.092
Dissatisfied	1,013.291	58.470	58.508	63.705	1.001	1.090
Very dissatisfied	314.068	32.856	32.858	36.135	1.000	1.100
Total	22,685.200	42.311	39.624	42.094	0.936	0.995
Median					1.000	1.092
Pay Grade						
E-4	479.836	38.213	38.266	41.078	1.001	1.075
E-5	4,592.477	103.630	103.552	99.824	0.999	0.963
E-6	9,315.055	135.076	135.018	131.009	1.000	0.970
E-7	7,400.000	126.401	126.383	111.853	1.000	0.885
E-8	747.268	49.409	49.412	46.387	1.000	0.939
E-9	101.168	16.608	16.608	15.171	1.000	0.913
Officers	0.000	0	0	0	-	-
Total	22,635.804	44.556	42.005	44.45	0.943	0.998
Median					1.000	0.951
Education						
Less than 12 years of school (no diploma)	55.471	12.166	12.166	11.386	1.000	0.936
GED or high school certificate	305.780	31.195	31.224	33.803	1.001	1.084
High School Diploma	3,737.716	94.077	94.069	98.912	1.000	1.051
High school diploma and some college, but did not graduate	11,627.437	140.149	140.031	158.442	0.999	1.131
Associates degree (E.G., AA, AS)	4,396.284	110.834	110.850	116.401	1.000	1.050

Table A-2.
A Comparison of Estimates Made From a File with Enlisted Treated as a Domain and a File Containing Enlisted Members Only (continued)

Item	Estimated Total	SUDAAN SEs		WesVar SEs	Ratio	
		Enlisted as a domain	Enlisted members only	Enlisted as a domain	SUDAAN enlisted only/ SUDAAN enlisted as a domain	WesVar domain /SUDAAN enlisted as a domain
Education (continued)						
Master's, Doctoral degree or professional school	191.671	25.498	25.499	23.763	1.000	0.932
Total	22,506.649	49.022	46.801	49.019	0.955	1.000
Median					1.000	1.050
Ethnicity						
Not Hispanic/Latino	20,163.166	95.218	94.405	97.653	0.991	1.026
Mexican/ Mexican American/ Chicano	1,022.049	55.079	55.117	60.42	1.001	1.097
Puerto Rican	687.873	50.410	50.43	55.395	1.000	1.099
Cuban	45.397	12.113	12.145	11.807	1.003	0.975
Other Spanish/ Hispanic/ Latino	532.701	41.996	42.019	41.628	1.001	0.991
Total	22,451.186	50.834	48.621	49.731	0.956	0.978
					1.001	1.026
Marital Status						
Now married	17434.689	122.108	121.654	122.524	0.996	1.003
Separated	983.492	57.056	57.079	52.664	1.000	0.923
Divorced	2579.959	89.092	89.121	92.433	1.000	1.038
Widowed	50.054	12.989	12.989	12.97	1.000	0.999
Never married	1610.991	70.935	70.958	75.511	1.000	1.065
Total	22,659.184	42.789	40.156	46.030	0.938	1.076
Median					1.000	1.003
Race/Ethnicity						
Hispanic American Indian/Alaskan Native	29.083	8.340	8.340	8.150	1.000	0.977
Hispanic Asian	22.751	8.142	8.142	7.873	1.000	0.967
Hispanic Black/African American	90.780	18.259	18.259	18.378	1.000	1.007
Hispanic Native Hawaiian/Other Pacific Islander	28.700	9.681	9.682	9.457	1.000	0.977
Hispanic White	1,434.434	68.717	68.752	64.499	1.001	0.939
Hispanic multiple race	55.308	14.105	14.105	14.025	1.000	0.994
Hispanic unknown race	650.571	45.298	45.329	46.074	1.001	1.017
Not Hispanic American Indian/Alaskan Native	204.844	25.904	25.921	27.338	1.001	1.055
Asian	466.726	38.653	38.689	38.794	1.001	1.004
Black/African American	4,527.688	110.915	110.936	123.852	1.000	1.117

Table A-2.
A Comparison of Estimates Made From a File with Enlisted Treated as a Domain and a File Containing Enlisted Members Only (continued)

Item	Estimated Total	SUDAAN SEs		WesVar SEs	Ratio	
		Enlisted as a domain	Enlisted members only	Enlisted as a domain	SUDAAN enlisted only/ SUDAAN enlisted as a domain	WesVar domain /SUDAAN enlisted as a domain
Race/Ethnicity (continued)						
Hawaiian/ Other Pacific Islander	89.960	18.495	18.497	17.305	1.000	0.936
White	14,478.654	135.277	134.93	153.003	0.997	1.131
American Indian/Alaskan Native and White	66.550	13.505	13.519	12.688	1.001	0.940
Asian and White	26.578	10.202	10.202	10.276	1.000	1.007
Black/African American and White	16.171	6.402	6.403	6.535	1.000	1.021
American Indian/Alaskan Native and Black	9.443	4.712	4.712	4.653	1.000	0.987
Non-Hispanic reporting 1+ race	76.305	17.113	17.115	18.141	1.000	1.060
Total	22,274.546	55.92	53.94	55.662	0.965	0.995
Median					1.000	1.004
Gender						
Male	20,805.225	85.118	84.304	103.007	0.990	1.210
Female	1,813.104	74.548	74.577	90.422	1.000	1.213
Total	22,618.329	45.587	43.097	45.084	0.945	0.989
Median					0.995	1.212
Service[1]						
Army	7,650.961	36.749	36.456	37.592	0.992	1.023
Navy	5,170.544	21.902	22.005	23.310	1.005	1.064
Marine Corps	2,606.997	7.846	7.846	7.812	1.000	0.996
Air Force	1,919.783	10.205	10.205	10.228	1.000	1.002
Coast Guard	377.922	4.817	4.817	5.403	1.000	1.122
Army Reserve	1,139.361	39.937	39.937	37.991	1.000	0.951
Army National Guard	2,601.014	21.289	20.801	21.991	0.977	1.033
Naval Reserve	527.551	19.483	16.428	15.237	0.843	0.782
Air Force Reserve	280.588	7.031	7.031	7.881	1.000	1.121
Air National Guard	449.083	9.304	9.304	10.273	1.000	1.104
Total	22,723.805	41.304	38.565	40.458	0.934	0.980
Median					1.000	1.028

[1] Service is based on self-reported Service in question 1.

Table A-2.
A Comparison of Estimates Made From a File with Enlisted Treated as a Domain and a File Containing Enlisted Members Only (continued)

Item	Estimated Total	SUDAAN SEs		SEs	Ratio	
		Enlisted as a domain	Enlisted members only	Enlisted as a domain	SUDAAN enlisted only/ SUDAAN enlisted as a domain	WesVar domain /SUDAAN enlisted as a domain
Stratum[2]						
Air Force	1,903.136	8.690	8.690	7.931	1.000	0.913
Air Force Reserve	305.591	3.129	3.129	3.701	1.000	1.183
Air National Guard	443.775	6.770	6.770	7.838	1.000	1.158
Army National Guard	2,632.606	18.102	17.509	19.138	0.967	1.057
Army Region 1	1,693.950	5.549	5.549	10.981	1.000	1.979
Army Region 3	1,491.000	0.000	0	8.861	-	-
Army Region 4	1,470.000	0.000	0	8.433	-	-
Army Region 5	1,113.986	4.921	4.921	7.570	1.000	1.538
Army Region 6	1,604.823	4.650	0	11.753	-	-
Army Reserve	1,489.432	15.843	15.843	16.389	1.000	1.034
Coast Guard	373.650	3.037	3.037	3.921	1.000	1.291
Marine Corps Reg. East	1,195.786	5.582	5.582	6.849	1.000	1.227
Marine Corps Reg. West	1,417.036	4.411	4.411	6.320	1.000	1.433
Navy Area 1	1,095.623	6.661	6.661	8.085	1.000	1.214
Navy Area 3	1,258.841	6.705	6.705	7.074	1.000	1.055
Navy Area 5	1,277.367	9.083	9.083	10.719	1.000	1.180
Navy Area 8	1,375.483	5.298	5.298	7.333	1.000	1.384
Navy Area Other	59.000	0.000	0	10.819	-	-
Naval Reserve	643.976	13.740	3.160	12.550	0.230	0.913
Total	22,845.062	35.141	31.834	37.029	0.906	1.054
Median					1.000	1.183
Median across all item categories					1.000	1.038
Median across totals					0.941	0.997

[2] The Service categories are based on coding on sample frame. These can differ from self-reported Service.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
xx-08-2002	Technical Report	November 2000 - August 2002

4. TITLE AND SUBTITLE
Weighting Report for the 2000 Military Recruiter Survey

5a. CONTRACT NUMBER
DASW01-98-M-2108

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Flores-Cervantes, I., Valliant, R., Zucker, A. B. and George, B. J.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Westat, Inc.
1650 Research Boulevard
Rockville, MD 20850

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Human Resources Activity
4040 North Fairfax Drive, Suite 200
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
DHRA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
2002-004

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In 1989, the first *Recruiter Survey* was administered to obtain baseline information regarding field recruiters' perceptions of issues related to recruiter quality of life. Since then, the *Recruiter Survey* has been administered in 1991, 1994, 1996, 1998 and 2000. The survey results provided are based on returns from active-duty Service production recruiters--those with at least one year of recruiting experience and assigned a goal/mission.

The *Weighting Report for the 2000 Military Recruiter Survey* contains information on weighting procedures for the survey dataset.

15. SUBJECT TERMS
military recruiting

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Andrea Zucker
U	U	U	SAR	40	19b. TELEPHONE NUMBER *(Include area code)* (703) 696-7178

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.